Geospatial Data on Your Own Terms

Bruce Godfrey, University of Idaho, Moscow, ID
Keywords: *GIS, geospatial web applications, geographic information services, geospatial web services, geospatial computation knowledge engine*
Citation: Godfrey, G. (2016). Geospatial Data on Your Own Terms. *PNLA Quarterly*, 80(2).

Abstract
Increasingly, geospatial data are being produced and consumed for many different purposes, ranging from urban planning to geographical analysis tasks related to sociological, historical, and literary studies. Everyone from students to scientists to professionals in many different scholarly and business disciplines has likely found—or is soon likely to find—themselves asking questions that can be answered using geospatial data. Consumers of geospatial data tend to want it on demand—wherever, whenever, and however they can get it—and of course they want this process made easy. This article highlights two recent ways in which librarians at the University of Idaho assisted researchers with the development of unique geospatial Web applications to advance research, to better communicate the results of research, and to empower researchers with the knowledge to take advantage of these technologies. Additionally, the author considers how emerging geospatial technologies may provide opportunities for libraries to build applications upon their geospatial repositories that will provide users with answers to their research questions.

Introduction

Until quite recently, the primary consumers and producers of geospatial data were only those professionals trained to make use of desktop geographic information systems (GIS) software. The producers of geospatial data would work with repository managers to transfer and make their data available, and the consumers would query geospatial catalogs in search of the datasets they needed to perform their work. Upon discovering the datasets they needed, these geospatial data consumers would download the items from the geospatial repository to their local machines and, finally, carry out their tasks in desktop GIS software.

As the GIS Librarian for the University of Idaho Library, I am quite familiar with the scenario described above, and I have seen first-hand that this process can be daunting for even the most confident students and researchers. As just one of many examples, in early 2015 I received a question from a research associate at the University of Idaho who was searching for GIS data concerning the total number of acres of rangeland in Idaho by county. Several datasets that would help answer her question reside in our geospatial repository; for example, the U.S. Bureau of Land Management, Idaho State Office, publishes grazing allotment information as a geospatial data set. The research associate was able to query the geospatial catalog and find items matching her keyword query, but she was then unsure how to work with those data (which were only available for download) in order to answer her initial research question. What she really needed, short of learning a desktop GIS software program, was a Web application to provide a way to explore those data to gain information and knowledge. Additionally, in the future, she may benefit from a geospatial computational knowledge engine that could potentially be created to provide an answer to her questions. She turned to me in the hope that I could somehow demystify these GIS data: she needed me to translate these data into a workable format to enable her to answer her question.

In the interest of assisting this patron in finding a timely answer, I downloaded the appropriate data sets, performed some geoprocessing in a desktop GIS program, produced a map, produced a table of needed data, and included citations to the dataset used. My skills in working for years with desktop GIS software enabled me to work through these steps, but the average academic library patron does not have this specialized training. While the solution in this case (connecting this researcher to GIS data in a workable, translatable form) involved the use of a desktop GIS program, a solution to these types of research questions could be more swiftly achieved through the development of geospatial Web applications that merely require a Web browser to access. This, I believe, is the future of geospatial data dissemination and knowledge acquisition: it must involve solutions that deliver geospatial data, information, and knowledge to users on their own time and on their own terms.

Nowadays, geospatial data are being utilized in many different ways, in many different public, private, and academic arenas; geospatial data are also being produced for and by many different audiences. Further, research questions are being posed in numerous disciplines that geospatial data can help answer. Those who are using geospatial data want to be able to derive information and knowledge from those data quickly, without being encumbered by technological obstacles.

In many cases, geospatial technologies have evolved to a point where users no longer need to, or have the desire to, download and store a local copy of data onto their desktop machine in order to gather the information and knowledge they require. That type of labor appears as an outmoded and rather cumbersome step that impedes users from accessing geospatial data that may be of great use to their research. These days, producers of geospatial data are also beginning to discover that it has become increasingly easier to publish these data they author as geospatial Web services so that desktop and mobile applications can link to the services. Furthermore, geospatial data producers are finding out that geospatial applications (which provide a context for users to better understand the data produced) are becoming ever easier to create. In short, as geospatial data consumers are coming to demand easier access to these data, these data are becoming easier than ever to share.

At the University of Idaho Library we have been embracing these changes. These are changes in geospatial technologies and, equally as important, changes in *expectations* about those technologies that represent a substantial opportunity for the academic library community to explore new ways to work for and with students and researchers. This article examines two recent examples where the author, in his role as GIS Librarian at the University of Idaho Library, assisted researchers with emerging geospatial technologies so that they could better communicate the results of their research. Additionally, the author contemplates how emerging technologies may provide research opportunities for librarians in creating applications that will provide users with answers to their research questions.

Background

Web-based GIS technologies have only been in existence since the mid-1990s (Dragićević, 2004). The Alexandria Digital Library (ADL) project, which commenced in 1994 and was managed out of the University of California Santa Barbara library, created a distributed system "…for collections of spatially indexed and graphical information" (Smith & Frew, 1995) and since then numerous libraries have created and maintained geospatial data catalogs over the years, using a variety of approaches (Kollen et al., 2013). These early systems have evolved into more sophisticated implementations that provide a range of functionality including preservation, curation, documentation, ingestion, visualization, and discovery services (Durante & Hardy, 2015; Wheeler & Benedict, 2015).

While libraries have provided various types of geographic information services for over two decades, recent advances in Web-based GIS functionality present new opportunities and challenges for the geographic information services offered by libraries. Geographic information services over the years have ranged from providing GIS computing labs, to providing assistance and consultation for locating data and working with GIS software, to managing geospatial data collections, and to providing instruction on the use of GIS technologies (Houser, 2006; Morris, 2006). Morris (2006) also discussed opportunities found and challenges experienced by libraries as they explored new roles related to geospatial Web services.

More recently, the creation of Web-based GIS applications has been exploding. Geospatial Web application development support is a new service component that could be added to the mix of services that academic libraries provide to their faculty and students. Requests for assistance with the development of Web-based geospatial applications are becoming more common; these patron requests are efforts to receive assistance in advancing geospatial-related research and to better communicate the results of geospatial-related research. In my own experience at the University of Idaho Library, I have also begun to see more frequent patron requests related to using cloud-based GIS platforms to enhance classroom instruction and research.

The development of Web-based GIS applications typically requires some programming skills such as JavaScript and HTML that utilize Application Programming Interfaces (APIs). Some cloud-based platforms have even begun to offer the creation of Web mapping applications using templates that require little or no programming knowledge. Cloud-based GIS platforms such as GIS Cloud (http://www.giscloud.com/), CartoDB (https://cartodb.com/cartodb-for/gis-online-software/), and ArcGIS Online (http://arcgis.com) place substantial application-development opportunities in the hands of the geospatial data producer. Given the increasing proliferation of use of these platforms, it is reasonable to contend that their use will only continue to grow. Geographic information services at academic libraries are well-positioned to respond to this growth with strategies in hand that can assist students and researchers in more easily utilizing geospatial data and in more effortlessly communicating the results of their work with these data.

Geospatial Web Services

The World Wide Web Consortium (W3C) defines a Web service as "…a software system designed to support interoperable machine-to-machine interaction over a network..." (W3C Working Group Note 2004). With increases in network connectivity and software enhancements from the early 2000s, geospatial Web services have become a viable way to share and access geographic information. There are several types of geospatial web services. Map services, arguably the most popular, make georeferenced map images available over the Internet: with such services, in many cases users no longer have to download and store a local copy of a dataset. Instead, geospatial Web services can be linked to from within applications. Applications or tools that consume data from Web services can be created to enable users to visualize and analyze data from desktop computer or mobile devices. Furthermore, geospatial Web services can provide data in preferred spatial reference systems or in specific formats and they can be utilized by developers to create applications and tools that quickly transform data into information and knowledge. Data that are accessible through Web services provides the foundation for geospatial Web application development: in short, making it easier to communicate information and knowledge. That same information and knowledge acquisition would take considerably longer, and would be significantly more difficult to obtain, if the only method of data access from a repository were to download the files.

Geospatial Web Applications

Geospatial Web applications provide a focused use, often of a map, and a context within which to use and understand geographic data. These applications enhance the explorability of data available from geospatial Web services and neatly package information for users to compre-hend.

Geospatial Web applications are developed using Application Programming Interfaces (API). Application Programming Interfaces are "…sets of requirements that govern how one application can talk to another [service or application]" (Proffitt, 2015). Developers of geospatial Web applications create standardized requests following the requirements of the API to create friendly user-interfaces for data. The user interacts with the application, the application 'talks' to the service, and the service 'talks' to the data. The consumer is rewarded with knowledge and information that the producer has prepared in a friendly package by creating a structure and meaning around the data. GIS librarians at academic libraries have an opportunity to work with students and researchers in developing geospatial Web applications and in utilizing cloud-based GIS platforms to help these patrons better advance their teaching and research and to better communicate the results of research.

Geospatial Web Application Development
What follows are two examples of recent instances in which the development of Web-based geospatial applications were used to meet the needs of researchers.

Example 1: Data Automation and Preparation
In the fall of 2014, a group of researchers assembled to kick-start a highly innovative, creative, and cross-institutional investigation into effective visualization strategies for ecosystem services in Idaho. These visualization strategies would allow researchers to better communicate the results of research to other researchers, stakeholders, and the public. Emphasis was on developing intuitive 3D visual interfaces to enable researchers, stakeholders, and the public to interactively view modeling products.

At the kick-off meeting, researchers noted their challenges with acquiring and preparing aerial imagery and topography data to be ingested into three-dimensional (3D) modeling software. The software requires data to be stored on local and/or network drives and will not ingest data from geospatial Web services. This group of researchers, the majority of whom weren't particularly familiar with geospatial data, described their difficulty navigating through the plethora of items available in numerous geospatial data catalogs. Data of interest, such as aerial imagery and digital elevation, were available, but a sobering reality emerged from the initial discussions: the user interface of the geospatial catalog (which we offered for discovery and access to data they needed) was of limited usefulness. The primary difficulties arose from the fact that these data were available in a variety of unfamiliar formats, in a variety of spatial reference systems, and with a variety of possible access methods. In short, the lack of normalization of these data was a consistent source of frustration for the researchers.

As the use of geospatial data continues to expand, users like these want to be able to use these data for their needs but they will not necessarily be experts in GIS. In other words, such researchers do not want to have to possess extensive knowledge about geospatial Web services, spatial reference systems, and other GIS concepts just to be able to use GIS data in their disciplines. User-centered applications need to be created to aid researchers in data preparation so that they can more swiftly get on with the business of doing their work.

As I listened to the group express their frustrations with the lack of access and ability to use these geospatial data, it was obvious that there was an opportunity for the library to provide an easier, more focused method of data acquisition and data preparation for researchers involved with this project. Clearly, a focused application (one that provided a focused use for data discoverable in the geospatial catalog) needed to be created. Even though the data for which they were searching were available as geospatial Web services, an additional step was necessary to prepare those data for use by these researchers. In short, and as stated at the meeting, they wanted to be able to "push a button and get the data" they needed: in this case a topography layer and an aerial imagery layer. I found myself asking, in this case, why couldn't our geospatial repository supply those data they needed to do their research on their terms? Why couldn't they just push a button and move onto focusing their valuable time and expertise on analysis? The automation of data preparation for ingestion into the downstream software needed to, and could be, streamlined; I realized that this streamlining could be accomplished by developing an in-house library geospatial Web service and a geospatial Web application. At this point, no service and software existed to specifically address the needs and desires of these researchers, but this dilemma quickly created an opportunity to present a solution. The dots were connected. I set out to get it done.

In order to meet these researchers' desires for easier data preparation and acquisition for this project, I designed a geospatial Web geoprocessing service and several different geospatial Web mapping applications. The development of the Web applications was performed using JavaScript and HTML, along with the ArcGIS API for JavaScript. The development for the geoprocessing service was achieved using Python. The geoprocessing services were published to an ArcGIS Server site at the University of Idaho Library. The geoprocessing service has many parameters, such as spatial reference system and image format hardcoded, but it could be adjusted to accept a variety of input parameters. The various components necessary to create this solution are as follows:

Web Application	**Geoprocessing Service**	**Web Services**	**Data Repository**
User interface that sends request to geoprocessing service	Processes request from web application and responds	Processes request from geoprocessing service and responds	Delivers data to web service
• HTML/JavaScript • Anyone could create their own application to send requests to the geoprocessing service	• Written in Python • Many options hard-coded. Optional parameters can be added.	• Current geoprocessing service operates on Esri Image Services • Could add capability for vector data.	• Data source for web services

The user interface examples created for this project were focused on three specific study areas in Idaho. However, the approach implemented has the flexibility to be utilized more broadly to prepare and deliver data to users for any study area for any project. Figure 1 shows the simplest interface where a user is simply able to push a button on a Web page and get data in the form of a downloadable file for a defined geographic location. This example serves to address the desire of the researchers to simply "push a button and get the data."

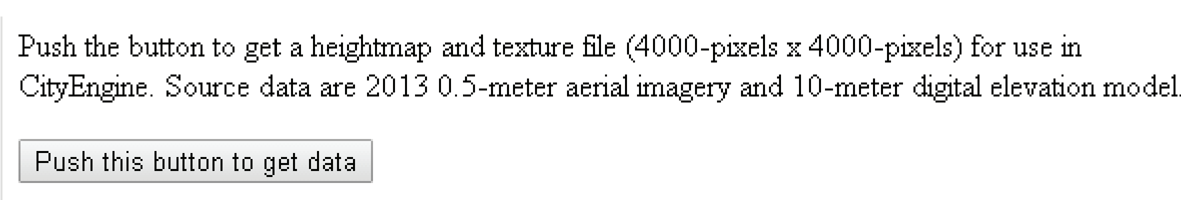


Figure 1 Provide the user with one button to get the data for a fixed geographic location.

Figure 2 shows an interface with multiple buttons that are configured to deliver data for various locations.

Push the button to get a heightmap and texture file (4000-pixels x 4000-pixels) for use in CityEngine. Source data are 2013 0.5-meter aerial imagery and 10-meter digital elevation model.

Push this button to get data for Fernan Lake

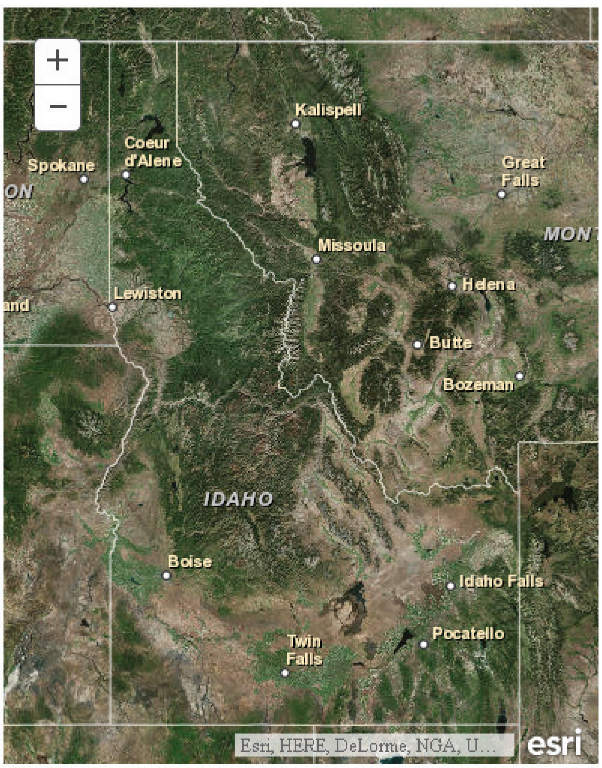
Push this button to get data for Boise

Push this button to get data for Pocatello

Figure 2 Provide the user with buttons for more than one geographic location.

Finally, Figure 3 introduces a web map that allows the user to choose the specific geographic location for which they would like data.

1. Zoom in and point-and-click on a location in Idaho.



2. Push this button to get data for the clicked location.

A heightmap and texture file (4000-pixels x 4000-pixels) for use in CityEngine will be returned for download. Source data are 2013 0.5-meter aerial imagery and 10-meter digital elevation model.

Figure 3 Web mapping application for selecting a geographic location.

Without question, having geospatial catalogs is helpful in many contexts for many users. But the opportunity to develop more focused user-centered/project-centered application interfaces that address the specific needs of researchers is upon us. In this case, a geospatial Web application and geospatial Web geoprocessing service needed to be draped over the existing infrastructure we had in place in order to add real value for the user. There is a real opportunity here for librarians to be active participants in advancing science and enabling research by understanding user needs and desires and by developing geospatial data applications and solutions.

Unfortunately, for this specific project, I was not able to provide the solution early enough to have a significant impact for the researchers. By the time I had developed the solution, they had the data they needed. However, the knowledge gained during this project positions our library to be more nimble, reactive, and potentially more importantly, to be more proactive in offering solutions when similar problems present themselves. The solutions we create in the future, based on the knowledge gained in this project, will allow researchers to be able to spend more time on analysis rather than data preparation and acquisition.

Example 2: Communicating the Results of Research

Librarians in the Data & Digital Services (DDS) Department of the University of Idaho Library also assisted researchers at the University of Idaho, Department of Civil Engineering in the fall of 2015, with communicating the results of current research they were conducting. The outputs from their research included a geospatial data set depicting normalized ground snow loads for Idaho. As the GIS Librarian, I worked with the Head of DDS to develop a geoprocessing model to generate the final output data set, to create documentation for the final output data set, to deposit the data set in the University of Idaho geospatial data repository, to publish the data set as a geospatial Web service, and to develop a Web mapping application to provide a context within which users could gain knowledge from those data. In short, the researchers were asking librarians in our department to assist them with advancing their research using geospatial technologies, documenting and curating their data, and with communicating the results of their research to the broader university and scholarly communities.

DDS was involved with this project from the outset and, as a department, was an active participant in the production of these scholarly data. Esri ArcGIS for Desktop was used to perform analysis and to generate the final data set. The software was further used to create a map containing the final data set which was subsequently published to an ArcGIS for Server site at the University of Idaho Library. Once the data set was finalized and made available as a geospatial Web service, development of a Web mapping application began. An HTML/JavaScript Web mapping application was developed using the Google Maps JavaScript Application Programming Interface (API) Version 3 and the ArcGIS Server Link for Google Maps API Version 3 (Figure 4). The map layer delivered from the Web service was overlaid on the Google Maps Terrain map type which we used as a basemap for geographic reference. An opacity slider—meant to allow users to adjust the transparency of the normalized ground snow load map layer—was added so that users could more clearly see the underlying terrain if they desired.

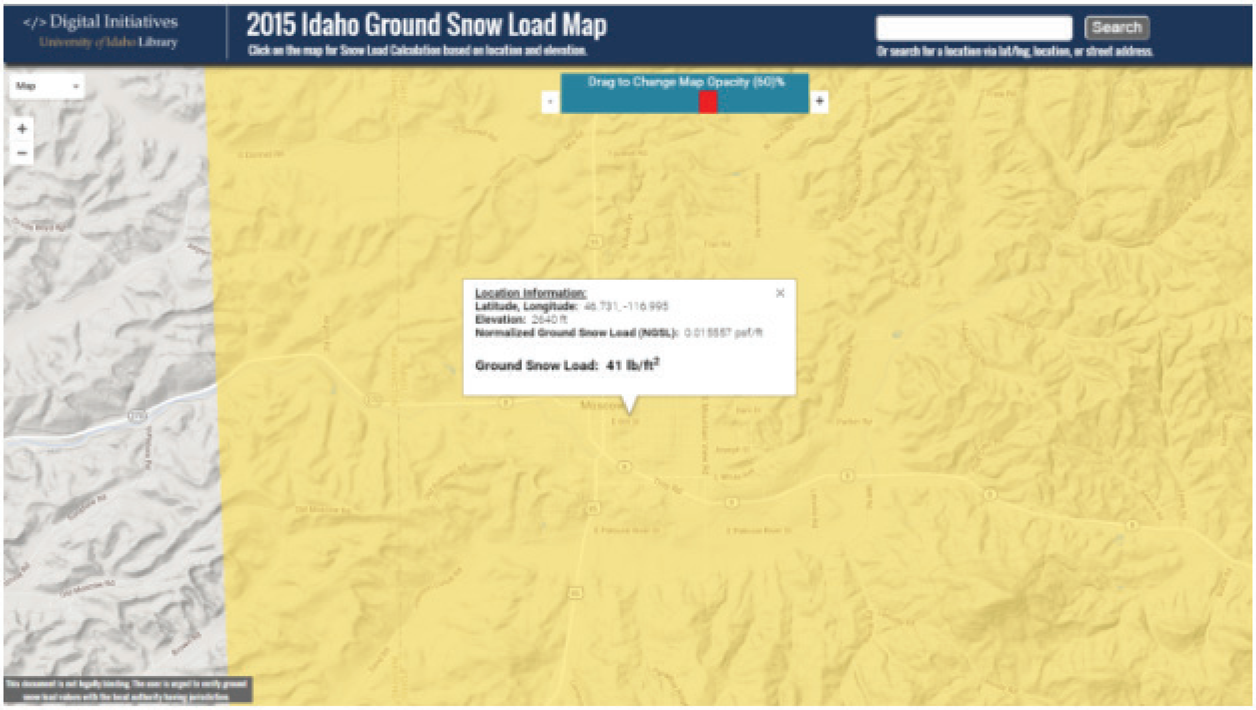


Figure 4 Web mapping application developed for Idaho Ground Snow Loads

Additional functionality was enabled to allow users to search for a geographic location by entering a town name, a latitude and longitude (in decimal degrees), or a street address. As a result of our efforts, users are now able to navigate around the map by zooming in/out and moving north/south/east/west and then clicking a location on the map to determine a ground snow load value for the selected location. The latitude and longitude, elevation, normalized ground snow load, and ground snow load are all displayed for the user for any selected location. The latitude and longitude, as well as the elevation value for any location are determined from the Google Maps API. The ground snow load is determined by multiplying the normalized ground snow load by the elevation in feet.

Having these data documented and placed in our geospatial repository is highly important, and having them available as a geospatial Web service is an added value. But the development of the Web mapping application was vital to communicate the results of this research and to make the information available to a broad audience. Additionally, the development of this additional Web mapping application was a vital service for the library to provide in order for us to be viewed as an active participant in communicating the results of scholarly output.

GIS in the Cloud

Over the past year I have been providing workshops on cloud-based GIS platforms such as ArcGIS Online. This has opened up several opportunities to discuss utilization of these technologies in humanities disciplines. For those interested in using these technologies, but who have little experience with the development of geospatial Web applications, these cloud-based platforms offer some exciting functionality. Namely, they offer templates that can be used to create geospatial Web applications with no programming required. The coming year looks bright for collaboration in both teaching and research, across many disciplines, using cloud-based GIS platforms.

Requests for assistance have focused on targeted training on the user interface and underlying functionality of the cloud-based platform. To date, researchers in several disciplines including English Literature, History, and Statistics have devoted time to training on how to use the ArcGIS Online platform for their research and teaching. Based on my experience over the last couple of years, use of these cloud-based GIS platforms is poised to grow significantly in the future.

Looking Ahead – Geospatial Computational Knowledge Engines?

If you query a search engine for the definition of a word, do you want the search engine to return links to dictionary Web sites? Not anymore; these days you expect the search engine to return the definition. In a similar fashion, why can't we query data in our academic geospatial repositories to answer questions directly? Can we begin to develop an application that sits atop the geospatial catalog and repository that will *answer* questions instead of providing links to ob-jects that contain the answers to questions?

I think libraries are poised to move beyond providing discovery of and access to objects that contain the answers to patrons' questions and, instead, to provide patrons with the answers to those questions directly. The opportunity to make this transformation widespread is especially present in the field of geographic information systems librarianship within academic research libraries. Confronting the challenge of how to provide patrons with tangible answers, rather than purely pointing them in the direction of answers, will inform library information science education and professional development by moving information organization toward the development of geospatial computational knowledge applications.

Currently, geospatial catalogs index content in their repositories, and in some cases other repositories, and provide users with a list of results of objects that match their text and/or spatial query. Merely providing patrons with access to objects that contain the information they are seeking is no longer enough; geospatial repositories and the catalogs that index their content must have a layer developed on top of them that functions as a computational knowledge engine (WolframAlpha, 2015) in order to take a transformative step to help patrons gain knowledge.

Case in point: In 1999, the Institute of Museum and Library Services awarded the University of Idaho Library a National Leadership Grant to develop a spatial and numeric geospatial data clearinghouse. INSIDE Idaho (http://insideidaho.org) was born from that award, and has endured for the past decade and a half, continuing to serve a wide variety of users with geospatial discovery and access services in the foreground while managing, curating, and preserving geospatial data in the background. The INSIDE Idaho repository contains within it a geospatial data set that depicts the average date when lilacs bloom in Idaho. This geospatial data set was produced as the output of research conducted at the University of Idaho and the University of Montana. Currently, if a user is looking to answer the question "When do lilacs bloom in Moscow, Idaho" that user can perform a text search, using the query string "lilac bloom," and they will be returned an item titled "Average Date when Lilacs Bloom in Idaho." The result sounds likely to address their information need. Within the catalog record they are presented several methods to access the data including as a downloadable GIS file, as a web service, or as a web application. Using one of these access methods they can, with a little time and effort, find the information for which they are searching—the answer to their question is in the data set.

Why then, since the answer they are seeking is in the data set that resides in the repository, can't a solution be developed that will return to the user the answer, a map, and a reference to the information source? That solution is the next evolution of geospatial repositories and catalogs; it is research that should be funded to take a transformative step toward knowledge discovery.

To realize the implementation of a geospatial computational knowledge engine, a sweeping and systematic collaboration across entities may be required. The end user will have to figure prominently into the usability design of the platform, the platform will need to be interoperable with existing systems, and professional development of librarians in enabling technologies will be required.

Conclusion

In each instance described above, librarians in the Data & Digital Services Department at the

University of Idaho Library worked to build a better bridge between researchers and geospatial data. Our work helped connect consumers to data more directly, more swiftly, and more efficiently. Our efforts in building these applications that provide a context for using these data allowed producers and consumers to quickly and easily interpret geospatial data and gain knowledge.

Some readers may balk at the general argument for a geospatial computation knowledge engine to increase ease and efficiency when it comes to academic investigation. (Here, I think of my own wife who is trained in humanities research, and I can hear her protest: "but research is not supposed to be easy or efficient! Think of how many happy, serendipitous accidents occur for researchers wandering the stacks of books, or searching numerous links in article databases!") But geospatial data delivery in its current state is limiting. There are too many insurmountable barriers that stand between 21st century researchers and those geospatial data they require. Users want software and solutions that are easy to use. And, in this case, I don't think we are losing anything of value in the research process by providing that to users.

And, finally, in no way am I arguing for the future irrelevance of GIS professionals in academic libraries; on the contrary, without GIS professionals, this kind of bridge-building could not be accomplished. I am merely pointing out, with the examples included here, that there are potentially better ways to connect patrons with data in our repositories and to provide answer to questions they may have. I think there are better ways on the horizon to deliver geospatial data to users on their own terms.

References

Dragićević, S. (2004). The potential of Web-based GIS. *Journal of Geographical Systems* 6(2), 79-81. doi:=10.1007/s10109-004-0133-4

Durante, K., & Hardy, D. (2015). Discovery, management, and preservation of geospatial data using hydra. *Journal of Map & Geography Libraries*, 11(2), 123-154. doi: 10.1080/15420353.2015.1041630

Houser, R. (2006). Building a library GIS service from the ground up. *Library Trends*, 55(2), 315-326.

Kollen, C., Dietz, C., Suh, J., & Lee, A. (2013). Geospatial data catalogs: Approaches by academic libraries. *Journal of Map & Geography Libraries*, 9(3), 276-295. doi: 10.1080/15420353.2013.820161

Morris, S. P. (2006). Geospatial web services and geoarchiving: New opportunities and challenges in geographic information services. *Library Trends*, 55(2), 285-303.

Proffitt, B. (2015). What APIs are and why they're important. *readwrite*. Retrieved from: http://readwrite.com/2013/09/19/api-defined

Smith, T. & Frew, J. (1995). Alexandria digital library. *Commun. ACM* 38(4), 61-62. doi: 10.1145/205323.205340

W3C Working Group Note. (2004). *Web services glossary*. Retrieved from: http://www.w3.org/TR/2004/NOTE-ws-gloss-20040211/#webservice

Wheeler, J. & Benedict, K. (2015) Functional requirements specification for archival asset management: Identification and integration of essential properties of services-oriented architecture products. *Journal of Map & Geography Libraries*, 11(2), 155-179. doi: 10.1080/15420353.2015.1035474

WolframAlpha (2015). *WolframAlpha*. Retrieved from: http://www.wolframalpha.com/

Bruce Godfrey is a GIS librarian at the University of Idaho where he has served in this capacity since October 2013. He holds a M.L.S from the University of North Texas, an M.S. in Environmental Science from the University of Idaho, and a B.A. in Environmental Science from the University of Virginia. Bruce can be contacted at bgodfrey@uidaho.edu.